

Computational Imaging and Vision

Computational Imaging and Vision

Managing Editor

MAX VIERGEVER

Utrecht University, Utrecht, The Netherlands

Series Editors

GUNILLA BORGEFORS, *Uppsala University, Uppsala, Sweden*

RAYMOND CHAN, *The Chinese University of Hong Kong, Hong Kong, China*

DANIEL CREMERS, *Technische Universität München, München, Germany*

RACHID DERICHE, *INRIA, Sophia Antipolis, France*

REINHARD KLETTE, *Auckland University of Technology, Auckland, New Zealand*

ALES LEONARDIS, *University of Ljubljana, Ljubljana, Slovenia*

STAN Z. LI, *CASIA, Beijing & Chinese Academy of Sciences, Beijing, China*

DIMITRIS N. METAXAS, *Rutgers University, New Brunswick, NJ, USA*

HEINZ-OTTO PEITGEN, *Universität Bremen, Bremen, Germany*

JOACHIM WEICKERT, *Universität des Saarlandes, Saarbrücken, Germany*

This comprehensive book series embraces state-of-the-art expository works and advanced research monographs on any aspect of this interdisciplinary field.

Topics covered by the series fall in the following four main categories:

- Imaging Systems and Image Processing
- Computer Vision and Image Understanding
- Visualization
- Applications of Imaging Technologies

Only monographs or multi-authored books that have a distinct subject area, that is where each chapter has been invited in order to fulfill this purpose, will be considered for the series.

Volume 45

More information about this series at <http://www.springer.com/series/5754>

Mahdi Rezaei • Reinhard Klette

Computer Vision for Driver Assistance

Simultaneous Traffic and Driver Monitoring



Springer

Mahdi Rezaei
Faculty of Computer and Information
Technology Engineering
Qazvin Branch
Islamic Azad University
Qazvin, Iran

Reinhard Klette
Department of Electrical and Electronic
Engineering
Auckland University of Technology
Auckland, New Zealand

ISSN 1381-6446
Computational Imaging and Vision
ISBN 978-3-319-50549-7 ISBN 978-3-319-50551-0 (eBook)
DOI 10.1007/978-3-319-50551-0

Library of Congress Control Number: 2016962187

Mathematics Subject Classification (2010): 68Txx, 68T45, 65Dxx, 65D19

© Springer International Publishing AG 2017, corrected publication 2018

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made.

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

The original version of this book was revised. An erratum to this book can be found at DOI [10.1007/978-3-319-50551-0_9](https://doi.org/10.1007/978-3-319-50551-0_9)

Preface

Currently, the automotive industry is experiencing three substantial changes of a revolutionary character, the replacement of internal combustion engines by electric motors, the integration of vehicle–vehicle and vehicle–infrastructure communication, and the rise of autonomous driving. The latter is still at a stage where drivers are expected to be in control of the vehicle at all times, but provided automated control features of the vehicle already enhance safety and driver comfort. We prefer to speak about *driver assistance* when addressing current developments towards autonomous driving. Driver-assistance features of vehicles are essentially based on data provided by various sensors such as radar, LIDAR, ultrasound, GPS, inertial measurement unit (IMU), or cameras. In this book, we discuss the use of cameras for driver assistance.

Computer vision-based driver assistance is an emerging technology, in both the automotive industry and academia. Despite the existence of some commercial safety systems such as night vision, adaptive cruise control, and lane departure warning systems, we are at the beginning of a long research pathway towards a future generation of intelligent vehicles.

Challenging lighting conditions in real-world driving scenarios, simultaneous monitoring of *driver vigilance* and *road hazards*, ensuring that the system responds in *real time*, and the legal requirements to comply with a high degree of *accuracy* are the main concerns for the developers of any advanced driver-assistance system (ADAS).

This book reviews relevant studies in the past decade as well as the state of the art in the field. The book also proposes various computer vision algorithms, techniques, and methodologies to address the aforementioned challenges, which are mainly suitable to be implemented for monocular vision.

Mobile devices (such as a mobile phone or a small-sized computing device) are often equipped with one or multiple cameras. If they also come with a stereocamera option, then this is typically with a small baseline (i.e., the distance between the two cameras), which supports stereo visualization, but not accurate stereo image analysis. Purpose-built processing units for driver monitoring or traffic observation, to be an add-on to already existing cars, also benefit from solutions for monocular

vision. However, our focus on monocular vision does not exclude the potential integration of the provided solutions into a multi-sensor environment, possibly also collecting data by using stereo vision, a laser range-finder, radar, ultrasound, GPS, an inertial measurement unit, or other sensors.

The discussion of monocular vision is also of academic interest, providing answers to the question of which tasks can be solved sufficiently by “using one eye only”, which defines an important research subject on its own.

The first part of the book focuses on monitoring driver vigilance, including classification, detection, and tracking of the driver’s *facial features*, i.e., eye status, head pose, yawning detection, and head nodding. The second part of the book mainly contributes to methods for road perception and road hazard monitoring, by introducing novel algorithms for *vehicle detection* and *distance estimation*. In the third part of the book, we simultaneously analyse the driver’s attention (in-cabin data) and road hazards (out-road data). We apply a *data fusion* approach on both data sources to measure the overall risk of driving conditions, to prevent or mitigate imminent crashes, and to assist a distracted driver in a timely and efficient manner.

For each part of our discussion, we present and analyse real-world experimental results, supported by benchmarks on a comprehensive range of datasets.

The book consists of eight chapters. Each chapter of the book addresses a major goal. All chapters follow a unique structure starting with an introduction and an overview of related work, followed by the proposed technique or algorithm(s). Then we compare the proposed methods against common existing techniques and the state of the art, followed by discussions on novelties and improvements based on empirical, analytical, or experimental results. The outline of the book is as follows:

Chapter 1: Vision-Based Driver-Assistance Systems

This chapter provides a general overview of tasks addressed when using cameras in modern cars.

Chapter 2: Driver–Environment Understanding

This chapter refines the general goals described in Chap. 1 for the specific context of interactions between driver monitoring and road environment; the chapter reviews related work published elsewhere.

Chapter 3: Computer Vision Basics

We briefly describe common computer vision concepts (theoretical and mathematical) which are of relevance for the following chapters in the book. This chapter includes a review on image notations, including integral images, colour conversion, line detection (Hough space), camera coordinates, and stereo vision analysis.

Chapter 4: Object Detection, Classification, and Tracking

This chapter provides the basics of classification and machine learning. It starts with a general introduction to object detection and then discusses supervised and unsupervised classification techniques. It ends with a brief outline of object tracking. This is material which is also of relevance for the following chapters.

Chapter 5: Driver Drowsiness Detection

We analyse the strengths and weaknesses of common face detectors such as LBP- and Haar-like-based detectors and their suitability for a DAS application. We discuss the causes of face detection failure in the context of driving conditions, followed

by our solutions to improve the results. The most important part of this chapter is the introduction of an *adaptive global* and *dynamic global* Haar-like classifier as a significant improvement on the standard Viola–Jones detector in both “training” and “application” phases. We also provide a Kalman filter-based tracking solution that can successfully track the driver’s drowsiness status, especially for open and closed eye detection at night, under sharp street lights, strong shades, strong backlights, and partial occlusions.

Chapter 6: Driver Inattention Detection

Following the initial facial features detected in the previous chapter, we continue by developing robust techniques in order to detect the driver’s direction of attention. The main aim of this chapter is to achieve a 6D head-pose estimation including roll, yaw, pitch, and (x, y) position of the head, at the time t . Having an accurate pose detection, as well as the eye-status detection discussed in Chap. 5, we are able to assess the driver’s level of fatigue or distraction.

Chapter 7: Vehicle Detection and Distance Estimation

In this chapter, we introduce an AGHaar-based vehicle detector augmented with multi-clue feature analysis and a Dempster–Shafer fusion. The chapter provides sufficient experiments to assure the robustness of the vehicle detection algorithm under different weather and challenging conditions.

Chapter 8: Fuzzy Fusion for Collision Avoidance

We can now define a risk level for every moment of driving, based on the online data acquired from Chaps. 5, 6, and 7 (i.e., driver’s state, detected vehicles on the road, and distance and angle of the ego-vehicle to the lead vehicles).

Within a fuzzy fusion platform, we correlate the in-vehicle data with the out-road hazards and identify the risk level of a potential crash driving condition based on the eight input parameters: yaw, roll, pitch (the driver’s direction of attention), eye status (open or closed), yawning, head nodding, distance to the detected vehicles in front of the ego-vehicle, and angle of the detected vehicle to the ego-vehicle.

The book briefly discusses stereo vision for driver assistance, but focuses in general on the analysis of monocular video data. We do not cover all the tasks which are possibly solvable by monocular vision only. For example, there is also traffic sign recognition, lane detection, and motion analysis (e.g., optic flow calculation), which are also areas experiencing substantial progress using monocular vision only. However, by discussing driver monitoring and examples of traffic monitoring, we present the required components that illustrate the important novel contribution of merging the analysis of both in-cabin data and out-road data.

The authors acknowledge the cooperation of colleagues, visitors, and students of the *Environment Perception and Driver Assistance* (.enpeda..) project, originally at the University of Auckland and now at Auckland University of Technology. We would especially like to thank Zahra Moayed and Noor Saleem for providing the text about the KLT tracker.

Contents

- 1 Vision-Based Driver-Assistance Systems** 1
 - 1.1 Driver-Assistance Towards Autonomous Driving..... 1
 - 1.2 Sensors 2
 - 1.3 Vision-Based Driver Assistance 4
 - 1.4 Safety and Comfort Functionalities 7
 - 1.5 VB-DAS Examples 8
 - 1.6 Current Developments..... 12
 - 1.7 Scope of the Book 15
- 2 Driver-Environment Understanding** 19
 - 2.1 Driver and Environment 19
 - 2.2 Driver Monitoring 20
 - 2.3 Basic Environment Monitoring 25
 - 2.4 Midlevel Environment Perception 30
- 3 Computer Vision Basics** 37
 - 3.1 Image Notations 37
 - 3.2 The Integral Image..... 39
 - 3.3 RGB to HSV Conversion..... 40
 - 3.4 Line Detection by Hough Transform 41
 - 3.5 Cameras 43
 - 3.6 Stereo Vision and Energy Optimization 44
 - 3.7 Stereo Matching 47
- 4 Object Detection, Classification, and Tracking** 51
 - 4.1 Object Detection and Classification 51
 - 4.2 Supervised Classification Techniques 53
 - 4.2.1 The Support Vector Machine 53
 - 4.2.2 The Histogram of Oriented Gradients 59
 - 4.2.3 Haar-Like Features 63

4.3	Unsupervised Classification Techniques	68
4.3.1	k -Means Clustering.....	68
4.3.2	Gaussian Mixture Models	72
4.4	Object Tracking.....	78
4.4.1	Mean Shift	80
4.4.2	Continuously Adaptive Mean Shift	84
4.4.3	The Kanade–Lucas–Tomasi (KLT) Tracker.....	85
4.4.4	Kalman Filter	89
5	Driver Drowsiness Detection.....	95
5.1	Introduction	95
5.2	Training Phase: The Dataset	97
5.3	Boosting Parameters.....	99
5.4	Application Phase: Brief Ideas	99
5.5	Adaptive Classifier.....	102
5.5.1	Failures Under Challenging Lighting Conditions.....	102
5.5.2	Hybrid Intensity Averaging	104
5.5.3	Parameter Adaptation	105
5.6	Tracking and Search Minimization	107
5.6.1	Tracking Considerations	107
5.6.2	Filter Modelling and Implementation	108
5.7	Phase-Preserving Denoising	110
5.8	Global Haar-Like Features	111
5.8.1	Global Features vs. Local Features	112
5.8.2	Dynamic Global Haar Features	114
5.9	Boosting Cascades with Local and Global Features	114
5.10	Experimental Results	115
5.11	Concluding Remarks	125
6	Driver Inattention Detection.....	127
6.1	Introduction	127
6.2	Asymmetric Appearance Models	129
6.2.1	Model Implementation	129
6.2.2	Asymmetric AAM.....	131
6.3	Driver’s Head-Pose and Gaze Estimation	133
6.3.1	Optimized 2D to 3D Pose Modelling.....	134
6.3.2	Face Registration by Fermat-Transform.....	136
6.4	Experimental Results	139
6.4.1	Pose Estimation	139
6.4.2	Yawning Detection and Head Nodding.....	139
6.5	Concluding Remarks	144
7	Vehicle Detection and Distance Estimation	147
7.1	Introduction	147
7.2	Overview of Methodology	149
7.3	Adaptive Global Haar Classifier	152

7.4	Line and Corner Features.....	155
7.4.1	Horizontal Edges	156
7.4.2	Feature-Point Detection	157
7.5	Detection Based on Taillights	159
7.5.1	Taillight Specifications: Discussion	159
7.5.2	Colour Spectrum Analysis	161
7.5.3	Taillight Segmentation	162
7.5.4	Taillight Pairing by Template Matching	163
7.5.5	Taillight Pairing by Virtual Symmetry Detection	165
7.6	Data Fusion and Temporal Information	168
7.7	Inter-vehicle Distance Estimation.....	171
7.8	Experimental Results	174
7.8.1	Evaluations of Distance Estimation	175
7.8.2	Evaluations of the Proposed Vehicle Detection	176
7.9	Concluding Remarks	187
8	Fuzzy Fusion for Collision Avoidance	189
8.1	Introduction	189
8.2	System Components	191
8.3	Fuzzifier and Membership Functions	192
8.4	Fuzzy Inference and Fusion Engine	195
8.4.1	Rule of Implication	196
8.4.2	Rule of Aggregation	196
8.5	Defuzzification.....	197
8.6	Experimental Results	197
8.7	Concluding Remarks	204
Erratum to: Computer Vision for Driver Assistance: Simultaneous Traffic and Driver Monitoring.....		E1
Bibliography		207
Index		221

Symbols

General Notation	Description
\mathbb{N}	Set of natural numbers
\mathbb{Q}	Set of rational numbers
\mathbb{R}	Set of real numbers
\mathbb{Z}	Set of integers
I	Image in spatial domain
N_{cols}	Image width in pixels
N_{rows}	Image height in pixels
x, y	Pixel coordinates (x, y) in a 2D image \mathbb{R}^2
u	Pixel intensity value
Ω	Image carrier, a set of all $N_{cols} \times N_{rows}$ pixel locations
X, Y, Z	Coordinates in a 3D space \mathbb{R}^3
p, q	Points in \mathbb{R}^2 , with coordinates x and y
P, Q, R	Points in \mathbb{R}^3 , with coordinates X, Y , and Z
W, W_p	Window in an image, windows with reference pixel location p
a, b, c	Real numbers
$\mathbf{a}, \mathbf{b}, \mathbf{c}$	Vectors
$\mathbf{A}, \mathbf{B}, \mathbf{C}$	Matrices
$\mathbb{A}, \mathbb{B}, \mathbb{C}$	Fuzzy sets
α, β, γ	Angles
$\mathcal{A}(\cdot)$	Area of a measurable set (function)
\angle	Angle sign
θ	Angular parameter in Hough space
ψ	Standard Haar-like feature
Φ	Global Haar-like feature
f	Camera focal length
h	Camera height from the ground
Θ	Camera viewing angle
H	Hue

S	Saturation
V	Value (in HSV colour space)
<i>t</i>	Time
t	Translation vector
τ	Threshold
R	Rotation matrix
ℓ	Length
ρ	Distance to origin in Hough space
ζ	Parity
\wedge	Logical AND
\vee	Logical OR
\cap	Intersections of sets
\cup	Union of sets
\times	Times (multiplication)
\oplus	Orthogonal sum
\otimes	Data fusion
\subset	Subset
\approx	Approximation
\in	Belonging to a set
ω	Training weights
δ	Saturation angle (in HSV colour space)
Γ	Coefficient factor
ε	Very small real greater than zero
Δ	Time or value difference between two variables
∇	Vector differential operator (e.g., for calculating the image gradient)
Π	Product
Σ	Summation